Self-teleportation and its application on LOCC estimation and other tasks

Keiji Matsumoto
National Institute of Informatics, 2-1-2, Hitotsubashi, Chiyoda-ku, 101-8430

February 2, 2008

Abstract

A way to characterize quantum nonlocality is to see difference in the figure of merit between LOCC optimal protocol and globally optimal protocol in doing certain task, e.g., state estimation, state discrimination, cloning and broadcasting. Especially, we focus on the case where n tensor of unknown states.

Our conclusion is that separable pure states are more non-local than entangled pure states. More specifically, the difference in the figure of the merit is exponentially small if the state is entangled, and the exponent is log of the largest Schmidt coefficient. On the other hand, in many cases, estimation of separable states by LOCC is worse than the global optimal estimate by $O\left(\frac{1}{n}\right)$.

To show that the gap is exponentially small for entangled states, we propose self-teleportation protocol as the key component of construct of LOCC protocols. Objective of the protocol is to transfer Alice's part of quantum information by LOCC, using intrinsic entanglement of $|\phi_\theta\rangle^{\otimes n}$ without using any extra resources. This protocol itself is of interest in its own right.

1 Introduction

A way to characterize quantum nonlocality is to see difference in the figure of merit between LOCC optimal protocol and globally optimal protocol in doing certain task, e.g., state estimation, state discrimination, cloning and broadcasting. Especially, we focus on the case where n tensor of unknown states: Alice and Bob starts from $|\phi_\theta\rangle^{\otimes n}$ with unknown θ and $|\phi_\theta\rangle$'s being a member of state family $\{|\phi_\theta\rangle\}_{\theta\in\Theta}$.

Our conclusion is that separable pure states are more non-local than entangled pure states. More specifically, the difference in the figure of the merit is exponentially small if $|\phi_\theta\rangle$ is entangled, and the exponent is log of the largest Schmidt coefficient. On the other hand, in many cases, estimation of separable states by LOCC is worse than the global optimal estimate by $O\left(\frac{1}{n}\right)$.

In the past, there had been many works on LOCC state detection. So far as I know of, there had been no substantial work about LOCC estimation of continuous unknown parameter. Past results on state detections are either about very specific case or the statement is rather weak. (see, for example [5], and references therein.) For example, conditions for the optimal measurement, the conditions for perfect detection, upperbound to the figure of the merit, and so on. In this work, by introducing asymptotic point of view, we can cover all the entangled pure states, and statement about the figure of the merit is had turned out to be the same as the global optimal measurement, which had been studied in detail.

To show that the gap is exponentially small for entangled states, we propose self-teleportation protocol as the key component of construct of LOCC protocols. Objective of the protocol is to transfer Alice's part of quantum information by LOCC, using intrinsic entanglement of $|\phi_\theta\rangle^{\otimes n}$ without using any extra resources. This protocol itself is of interest in its own right.

To show the gap is large for separable pure states, we develop general theory of LOCC estimation of tensor product states, and show the sufficient conditions for $O\left(\frac{1}{n}\right)$ gap to be observed between LOCC optimum and the global optimum.

2 Self-teleportation

2.1 objective

Suppose Alice and Bob share n copies of a unknown $d\times d$ bipartite pure state $|\phi\rangle_{AB}$. The objective is to transfer the Alice's quantum information to Bob's local space by LOCC without extra resources, with exponentially high fidelity:

$$|\phi\rangle\langle\phi|^{\otimes n}\in\mathcal{S}\left(\mathcal{H}_A^{\otimes n}\otimes\mathcal{H}_B^{\otimes n}\right)\underset{\text{LOCC}}{\longrightarrow}\rho_\phi^n\in\mathcal{S}\left(\mathcal{H}_B^{\otimes n}\otimes\mathcal{H}_{B'}^{\otimes n}\right),$$

and

$$\left\langle\phi^{\otimes n}\right|\rho_\phi^n\left|\phi^{\otimes n}\right\rangle\geq 1-O\left(\left(p_1^\phi\right)^n\right),$$

where $\mathcal{H}_A^{\otimes n}$ and $\mathcal{H}_B^{\otimes n}\otimes\mathcal{H}_{B'}^{\otimes n}$ is at Alice and Bob's side, respectively, and $p_1^\phi\geq p_2^\phi\geq\cdots\geq p_d^\phi$ are the Schmidt coefficient of $|\phi\rangle$.

In case $|\phi\rangle$ is entangled, $p_1^\phi<1$, and the fidelity is exponentially close to 1, while it equals 0 for a tensor product state.

2.2 A standard form of an ensemble of identical bipartite pure states

Note $|\phi\rangle^{\otimes n}$ is invariant by the reordering of copies, or the action of the symmetric group S_n. Action of the symmetric group occurs a decomposition of the tensored space $\mathcal{H}^{\otimes n}$ [14],

$$\mathcal{H}^{\otimes n} = \bigoplus_{\lambda:\lambda\vdash n} \mathcal{W}_\lambda, \; \mathcal{W}_\lambda := \mathcal{U}_\lambda \otimes \mathcal{V}_\lambda,$$

where $\mathcal{U}_\lambda$ and $\mathcal{V}_\lambda$ is an irreducible space of the tensor representation of $\mathrm{SU}(d)$, and the representation of S_n, respectively, and $\lambda = (\lambda_1, \ldots, \lambda_d)$ ($\lambda_i \geq \lambda_{i+1} \geq 0$, $\sum_{i=1}^{d} \lambda_i = n$) is called *Young index*, which $\mathcal{U}_\lambda$ and $\mathcal{V}_\lambda$ uniquely corresponds to. We denote by $\mathcal{U}_{\lambda,A}$, $\mathcal{V}_{\lambda,A}$, and $\mathcal{U}_{\lambda,B}$, $\mathcal{V}_{\lambda,B}$ the irreducible component of $\mathcal{H}_A^{\otimes n}$ and $\mathcal{H}_B^{\otimes n}$, respectively. Also, $\mathcal{W}_{\lambda,A}$ and $\mathcal{W}_{\lambda,B}$ are defined analogously.

In terms of this decomposition, $|\phi\rangle^{\otimes n}$ can be written as

$$|\phi\rangle^{\otimes n} = \bigoplus_{\lambda:\lambda\vdash n} a_\lambda \left|\phi_\lambda\right\rangle \left|\Phi_\lambda\right\rangle, \tag{1}$$

where $|\phi_\lambda\rangle \in \mathcal{U}_{\lambda,A} \otimes \mathcal{U}_{\lambda,B}$, and $|\Phi_\lambda\rangle \in \mathcal{V}_{\lambda,A} \otimes \mathcal{V}_{\lambda,B}$. While a_λ and $|\phi_\lambda\rangle$ are dependent on $|\phi\rangle$, $|\Phi_\lambda\rangle$ is a maximally entangled state which does not depend on $|\phi\rangle$,

$$\left|\Phi_\lambda\right\rangle := \frac{1}{\sqrt{d_\lambda}} \sum_{i=1}^{d_\lambda} \left|f_{A,i}^\lambda\right\rangle \left|f_{B,i}^\lambda\right\rangle.$$

Here, $\left\{\left|f_{A,i}^\lambda\right\rangle\right\}$ and $\left\{\left|f_{B,i}^\lambda\right\rangle\right\}$ is a CONS of $\mathcal{V}_{\lambda,A}$ and $\mathcal{V}_{\lambda,B}$, respectively, and $d_\lambda := \dim \mathcal{V}_\lambda$.

2.3 Protocol and performance

Rough description of our protocol is as follows. Alice teleports her part of $|\phi_\lambda\rangle$ using $|\Phi_\lambda\rangle$. $|\Phi_\lambda\rangle$'s will be gone, but can be reconstructed by Bob since it is independent of $|\phi\rangle$. Schmidt rank, or equivalently d_λ is small than $\dim \mathcal{U}_\lambda$ for most of λ, and our fidelity of success is exponentially close to 1.

Note that the following protocol does not work: Upon measurement of $\mathcal{W}_{\lambda,A} \otimes \mathcal{W}_{\lambda,B}$ (In the paper, the projector will be denoted by the same symbol as its support.), teleport $|\phi_\lambda\rangle$ using $|\Phi_\lambda\rangle$. In such a protocol, coherence between the subspaces corresponding to different values of λ will be destroyed. To keep the coherence, we use measurement which does not distinguish the Young index λ. Consider the following operators.

$$A_{\{U_\lambda\}} := \bigoplus_{\lambda\in \mathrm{A}_n} \sqrt{d_\lambda} \sum_{i=1}^{\dim \mathcal{U}_\lambda} \left\langle e_{A,i}^\lambda\right| \left\langle f_{A,i}^\lambda\right| U_\lambda^\dagger.$$

Here, $\mathrm{A}_n := \{\lambda : \dim \mathcal{U}_\lambda \leq d_\lambda\}$, $\left\{\left|e_{A,i}^\lambda\right\rangle\right\}$ is a CONS of $\mathcal{U}_{A,\lambda}$, and U_λ is runs all over the elements of $\mathrm{U}\left(\mathcal{V}_{A,\lambda}\right)$ (NB not $\mathrm{SU}\left(\mathcal{V}_{A,\lambda}\right)$). Observe that $\int U_\lambda \mathrm{d}U_\lambda = 0$ where $\mathrm{d}U_\lambda$ is an invariant measure in $\mathrm{U}(\mathcal{U}_\lambda)$ with the normalization $\int \mathrm{d}U_\lambda = 1$, since $-U_\lambda$ is also in $\mathrm{U}(\mathcal{V}_\lambda)$. Due to this and the Shur's lemma 5 and 4, we have

$$\int A_{\{U_\lambda\}}^\dagger A_{\{U_\lambda\}} \prod_{\lambda\in \mathrm{A}_n} \mathrm{d}U_\lambda = \bigoplus_{\lambda\in \mathrm{A}_n} \mathcal{W}_{\lambda,A} \otimes \mathcal{W}_{\lambda,B}.$$

(I) Alice and Bob project the state onto the subspace $\bigoplus_{\lambda\in \mathrm{A}_n} \mathcal{W}_{\lambda,A}$ and $\bigoplus_{\lambda\in \mathrm{A}_n} \mathcal{W}_{\lambda,B}$ If both of them succeed, they proceed.

In fact, from (1), if one of them succeeds, both of them succeeds.

The success probability equals $\sum_{\lambda\in \mathrm{A}_n} a_\lambda^2$.

(II) Alice Apply the measurement corresponding to $\left\{A_{\{U_\lambda\}}\right\}$, and send the measurement outcome $\{U_\lambda\}_{\lambda:\lambda\vdash n}$ to Bob. After this, Bob has the state

$$\bigoplus_{\lambda\in \mathrm{A}_n} a_\lambda \sum_{i=1}^{\dim \mathcal{U}_\lambda} \left\langle e_{A,i}^\lambda \middle| \phi_\lambda\right\rangle \overline{U_\lambda} \left|f_{B,i}^\lambda\right\rangle.$$

(III) Bob applies the recovery operation $\bigoplus_{\lambda\in \mathrm{A}_n} \mathbf{1}_{\mathcal{U}_B} \otimes U_\lambda^T$, to obtain $\bigoplus_{\lambda\in \mathrm{A}_n} a_\lambda |\phi_\lambda\rangle$, where $|\phi_\lambda\rangle$ is in $\mathcal{U}_{\lambda,B} \otimes \mathcal{U}_{\lambda,B'}$. Finally, he reconstruct $|\Phi_\lambda\rangle$ in $\mathcal{V}_{\lambda,B} \otimes \mathcal{V}_{\lambda,B'}$ to obtain

$$\bigoplus_{\lambda\in \mathrm{A}_n} a_\lambda \left|\phi_\lambda\right\rangle \left|\Phi_\lambda\right\rangle. \tag{2}$$

The fidelity between the final state and $|\phi\rangle^{\otimes n}$, in average, equals $\sum_{\lambda\in \mathrm{A}_n} a_\lambda^2$.

Using , we can evaluate $\sum_{\lambda\in \mathrm{A}_n} a_\lambda^2$:

$$\sum_{\lambda\in \mathrm{A}_n} a_\lambda^2 \geq 1 - \frac{d\,(2d-3)!}{(d-2)!\,(d-1)!} (n+1)^{\frac{d(d+1)}{2}} \left(p_1^\phi\right)^n$$

3 Application to LOCC state estimation

Suppose Alice and Bob share $n$copies of $|\phi_\theta\rangle$, with unknown continuous parameter $\theta \in \Theta \subset \mathbb{R}^D$. Their task is to estimate θ as accurately as possible by LOCC. Our interest is the difference in the figure of merit between LOCC optimal and the global optimal.

Consider the composition of the globally optimal protocol after the globally optimal measurement, we observe the figure of the merit differs only by exponentially small amount, if $|\phi_\theta\rangle$ is entangled.

More specifically, there are two kinds of error measures which commonly used. First one is of the form

$$\mathrm{E}\, dist\left(\theta, \hat{\theta}_n\right), \tag{3}$$

where $dist\,(,)$ is a smooth distance function, and E stands for the expectation about the random variable $\hat{\theta}_n$. The second one is

$$\beta_{\epsilon,\theta}^n := \Pr\left\{\left\|\hat{\theta}_n - \theta\right\| > \epsilon\right\}. \tag{4}$$

The first leading term of (3) is, if $dist\left(,\right)$ is smooth enough, $O\left(\frac{1}{n}\right)$. The second leading term is, looking back classical case, $\Omega\left(n^{-2}\right)$, and the third leading term is $\Omega\left(n^{-\frac{5}{2}}\right)$, and so on. Therefore, our LOCC measurement strategy is as good as the given protocol up to the higher order terms. The error measure (4) behaves as follows:

$$\beta_{\epsilon,\theta}^{n}\sim e^{-n\,O(\epsilon)}.$$

Especially, they are interested in the case that ϵ is small. If ϵ is small enough for the exponent of of $\beta_{\epsilon,\theta}^{n}$ to be smaller than $\log p_1^{\phi}$, our protocol is optimal.

4 Other applications

4.1 State detection

Suppose θ takes discrete values, and our aim is to estimate the parameter θ. Such a problem is called 'state detection'. Since $|\phi_\theta\rangle$ and $|\phi_{\theta'}\rangle$ are distant by some constant for all θ and θ', the error probability drops exponentially as n increases. If its exponent is smaller than the one of the RHS of $\log p_1^{\phi_\theta}$, our LOCC protocol will be as good as the given protocol in the main part of the error. For example, we discuss the sum of all the possible errors, given the candidates of the true state $|\phi_\theta\rangle$ $(\theta=1,\cdots,M)$. The error exponent cannot be smaller than the one for the detection problem $|\phi_\theta\rangle$ versus $|\phi_{\theta'}\rangle$. Therefore, $\max_{\theta\neq\theta'}\left|\langle\phi_\theta|\,\phi_{\theta'}\rangle\right|^2\geq\max_\theta p_1^{\phi_\theta}$ is the sufficient condition for the optimal LOCC measurement to achieve the global optimum. This condition holds if $|\phi_\theta\rangle$ are distant from tensor product states.

4.2 Cloning, broadcast

Suppose θ is continuous, and the family $\{|\phi_\theta\rangle\}$ is the totality of the pure states in d-dimensional Hilbert space. Now, our task is to make m copies of $|\phi_\theta\rangle_{BB'}$ from n copies of $|\phi_\theta\rangle_{AB}$. The optimal Fidelity of n to m cloning and broadcast with global operations is $(r)^{d-1}+O\left(\left(\frac{1}{n}\right)\right)$ and $1+\frac{r-1}{rn}+O\left(\frac{1}{n^2}\right)$, with $m=rn$ [13], [9]. Seeing these terms, we can observe the first several terms are obviously larger than exponential order. Hence, our teleportation scheme does not degrade these terms.

5 Asymptotic theory estimation of pure separable states

If $|\phi_\theta\rangle = |\phi_{A,\theta}\rangle\,|\phi_{B,\theta}\rangle$, we cannot use the self-teleportation. Hence, this case has to studied separately. In this section, we show some sufficient conditions for $O(1/n)$ gap to exist.

5.1 Asymptotic theory of estimation of pure states

Given a unknown system, a statistician, assuming that the state corresponds to a member of a family $\left\{\rho_\theta\,;\theta\in\Theta\subset\mathbb{R}^d\right\}$ of density matrix, applies a measurement M, obtain data, and calculate the estimate $\hat{\theta}$ of θ based on the measurement result. In asymptotic setting, we assume that $\rho_\theta^{\otimes n}$ is given, and the measurement M^n may act correctively on $\rho_\theta^{\otimes n}$. A common error measure is (3), where $dist\left(\cdot,\cdot\right)$ is smooth enough:

$$dist\left(\rho_\theta,\rho_{\theta+\mathrm{d}\,\theta}\right)=\sum_{i,j}G_{ij}\mathrm{d}\theta^i\mathrm{d}\theta^j+o\left(\mathrm{d}\theta\right)^2.$$

With such natural setting, the first leading term of (3) is $O\left(1/n\right)$, and our interest is to minimize the coefficient.

It is known that the optimal coefficient writes

$$\inf_{\{M^n\}}\varlimsup_{n\to\infty}n\mathrm{Tr}\,G\left(J_\theta^{M^n}\right)^{-1}. \tag{5}$$

Here $G=[G_{i,j}]$ is a real positive symmetric matrix, Tr is trace over $\mathbb{R}^d$, and $J_\theta^{M^n}$ is the *Fisher information matrix* of $p_\theta^{M^n}(x):=\mathrm{tr}\,\rho_\theta^{\otimes n}M_x$, which is defined by

$$\left[J_\theta^{M^n}\right]_{i,j}:=\sum_x p_\theta^{M^n}(x)\left(\frac{\partial\log p_\theta^{M^n}(x)}{\partial\theta^i}\,\frac{\partial\log p_\theta^{M^n}(x)}{\partial\theta^j}\right). \tag{6}$$

As for pure state models, we can explicitly characterize the Fisher information of the asymptotically optimal measurement [10]. First notable fact is that the collective measurement is not effective:

$$\min_{M^n}n\mathrm{Tr}\,G\left(J_\theta^{M^n}\right)^{-1}=\min_M\mathrm{Tr}\,G\left(J_\theta^M\right)^{-1},$$

where M in the RHS is a measurement acting on the single copy $|\phi\rangle$. In other words, defining

$$\mathfrak{J}_\theta^n:=\left\{\frac{1}{n}J_\theta^{M^n}\,;M^n:\text{ arbitrary collective meas.}\right\}.$$

we have

$$\mathfrak{J}_\theta^n=\mathfrak{J}_\theta^1.$$

Below, more quantitative results are in order. Define the matrix

$$J_{\theta,i,j}^S:=\Re\left\langle l_{\theta,i}\,|l_{\theta,j}\right\rangle$$

and

$$\tilde{J}_{\theta,i,j}:=\Im\left\langle l_{\theta,i}\,|l_{\theta,j}\right\rangle,$$

where

$$|l_{\theta,i}\rangle:=\frac{1}{2}\left\{\frac{\partial\,|\phi_\theta\rangle}{\partial\theta^i}-\langle\phi_\theta|\,\frac{\partial}{\partial\theta^i}\,|\phi_\theta\rangle\,|\phi_\theta\rangle\right\}.$$

These quantities are known to have tight connection with Berry's geometric phase. Namely, both the line integral of $|l_{\theta,i}\rangle$ along a closed curve and the area

integral of $\tilde{J}_{\theta,i,j}$ over the surface enclosed by the curve equals Berry phase.

The eigenvalue of $J_\theta^{S-1}\tilde{J}_\theta$ is in the form of $\pm\beta_{\theta,j}$ with

$$0 \le \beta_{\theta,j} \le 1.$$

$\beta_{\theta,j}$ are invariant by the change of the coordinate, and is closely related to a natural complex structure of the space of pure states. Indeed, $\arccos\beta_{\theta,j}$ are called multiple Kaehler angles. $\arccos\beta_{\theta,j} = 0$ $(\forall j)$ means that the state family is a complex submanifold.

Below, we mainly treat the case $\dim\theta = 2$. In this case, we denote $\beta_{\theta,1}$ by β_θ. Suppose that $dist(\cdot,\cdot)$ is the Bure's distance,

$$\begin{aligned} dist\left(\left|\phi_\theta\right\rangle, \left|\phi_{\theta+\mathrm{d}\theta}\right\rangle\right) &= \sqrt{1-\left|\left\langle\phi_\theta \,\middle|\phi_{\theta+\mathrm{d}\theta}\right\rangle\right|^2} \\ &= \frac{1}{2}\sum_{i,j} J^S_{\theta,i,j}\mathrm{d}\theta^i\mathrm{d}\theta^j + o\left(\mathrm{d}\theta\right)^2. \end{aligned}$$

If we use this measure of the error,

$$\begin{aligned} &\inf_{\{M^n\}}\varlimsup_{n\to\infty} n\sqrt{1-\left|\left\langle\phi_\theta \,\middle|\phi_{\hat{\theta}^n}\right\rangle\right|^2} \\ &= \min_M \operatorname{Tr} J^S_\theta\left(J^M_\theta\right)^{-1} \\ &= \sum_j \frac{4}{1+\sqrt{1-\beta^2_{\theta,j}}}. \end{aligned} \tag{7}$$

Hence, the error is monotone increasing in $\beta_{\theta,j}$.

5.2 The asymptotic estimation of tensor product pure state by LOCC

Suppose

$$\rho_\theta = \rho_{A,\theta}\otimes\rho_{B,\theta},$$

with $\rho_{A,\theta}$ and $\rho_{B,\theta}$ is supported on $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, and define $J_\theta^{M^{A,n}}$ by replacing M^n and $\rho_\theta^{\otimes n}$ in (6) by $M^{A,n}$ and $\rho^{\otimes n}_{A,\theta}$, respectively. $J_\theta^{M^{B,n}}$ is also similarly defined. As in appendix, slight modification of the proof of lemma 1 in [4] leads to:

Lemma 1 *If M^n is a LOCC measurement, there is a measurement $M^{A,n}$ and $M^{B,n}$ acting on $\mathcal{H}_A^{\otimes n}$ and $\mathcal{H}_B^{\otimes n}$, respectively, and satisfies*

$$J_\theta^{M^n} = J_\theta^{M^{A,n}} + J_\theta^{M^{B,n}}.$$

Here $J_\theta^{M^{A,n}}$ is defined by replacing M^n and $\rho_\theta^{\otimes n}$ in (6) by $M^{A,n}$ and $\rho^{\otimes n}_{A,\theta}$, respectively. $J_\theta^{M^{B,n}}$ is also similarly defined.

The lemma will be proved in subsection 5.3. This lemma means that LOCC measurement M^n can be replaced by the local measurements $M^{A,n}$ and $M^{B,n}$, so far as the Fisher information is concerned.

Define

$$\begin{aligned} &\mathfrak{J}_\theta^{LOCC,n} \\ &:= \left\{ \frac{1}{n}J_\theta^{M^n};\ \begin{array}{c} M^n\text{:LOCC in A-B split,} \\ \text{collective over } n \text{ copies} \end{array} \right\}, \end{aligned}$$

this lemma can be expressed as

$$\begin{aligned} &\mathfrak{J}_\theta^{LOCC,n} \\ &= \left\{ J\,;\ J = J^A + J^B, J^A \in \mathfrak{J}_\theta^{A,n}, J^B \in \mathfrak{J}_\theta^{B,n} \right\}. \end{aligned}$$

In pure state family case, as we seen above, $\mathfrak{J}_\theta^n = \mathfrak{J}_\theta^1$. Hence, combined with lemma 1, we can conclude that $\mathfrak{J}_\theta^{LOCC,n} = \mathfrak{J}_\theta^{LOCC,1}$, meaning that being collective over the copies is not useful also LOCC case, either.

Define $\left|l^A_{\theta,i}\right\rangle$, $J_\theta^{S,A}$, $\tilde{J}^A_\theta$, $\beta^A_{\theta,j}$ and $\left|l^B_{\theta,i}\right\rangle$, $J_\theta^{S,B}$, $\tilde{J}^B_\theta$, $\beta^B_{\theta,j}$ by replacing $\left|\phi_\theta\right\rangle$ by $\left|\phi^A_\theta\right\rangle$ and $\left|\phi^B_\theta\right\rangle$, respectively. Observe that

$$\begin{aligned} \left|l_{\theta,i}\right\rangle &= \left|l^A_{\theta,i}\right\rangle\left|\phi^B_\theta\right\rangle + \left|\phi^A_\theta\right\rangle\left|l^B_{\theta,i}\right\rangle, \\ J^S_\theta &= J^{S,A}_\theta + J^{S,B}_\theta, \qquad (8) \\ \tilde{J}_\theta &= \tilde{J}^A_\theta + \tilde{J}^B_\theta. \qquad (9) \end{aligned}$$

Further, we suppose $\dim\theta = 2$ and

$$J^{S,A}_\theta = aA,\ J^{S,B}_\theta = bA$$

with $A = A^T \ge 0$. This means with proper coordinate system, $J^{S,A}_\theta = a\mathbf{1}$ and $J^{S,B}_\theta = b\mathbf{1}$. Hence, we have

$$\beta_\theta = \frac{a\beta^A_\theta \pm b\beta^B_\theta}{a+b}, \tag{10}$$

and

$$\begin{aligned} &\max_{J\in\mathfrak{J}^1_\theta}\operatorname{Tr} J_\theta^{S-1}J \\ &= 1+\sqrt{1-\left(\frac{a\beta^A_\theta \pm b\beta^B_\theta}{a+b}\right)^2} \\ &\ge \frac{1}{a+b}\left(\begin{array}{c} a\left(1+\sqrt{1-\left(\beta^A_\theta\right)^2}\right) \\ +b\left(1+\sqrt{1-\left(\beta^B_\theta\right)^2}\right) \end{array} \right), \end{aligned}$$

where the identity holds if and only if $\beta^A_\theta = \beta^B_\theta$ and the $+$-sign in (10) is the case. On the other hand,

$$\begin{aligned} &\max_{J\in\mathfrak{J}^{LOCC,1}_\theta}\operatorname{Tr} J_\theta^{S-1}J \\ &= \max_{J\in\mathfrak{J}^{A,1}_\theta}\operatorname{Tr} J_\theta^{S-1}J + \max_{J\in\mathfrak{J}^{B,1}_\theta}\operatorname{Tr} J_\theta^{S-1}J \\ &= \frac{a}{a+b}\max_{J\in\mathfrak{J}^{A,1}_\theta}\operatorname{Tr}\left(J^{S,A}_\theta\right)^{-1}J + \frac{b}{a+b}\max_{J\in\mathfrak{J}^{B,1}_\theta}\operatorname{Tr}\left(J^{S,B}_\theta\right)^{-1}J \\ &= \frac{1}{a+b}\left\{ \begin{array}{c} a\left(1+\sqrt{1-\left(\beta^A_\theta\right)^2}\right) \\ +b\left(1+\sqrt{1-\left(\beta^B_\theta\right)^2}\right) \end{array} \right\}. \end{aligned}$$

Therefore, we have

$$\max_{J\in\mathfrak{J}^1_\theta}\operatorname{Tr} J_\theta^{S-1}J \ge \max_{J\in\mathfrak{J}^{LOCC,1}_\theta}\operatorname{Tr} J_\theta^{S-1}J.$$

Due to [10], the maximum of $\operatorname{Tr} J_\theta^{S-1}J$ and the minimum of $\operatorname{Tr} J_\theta^{S}J^{-1}$is achieved by the same matrix, a constant multiple of J^S_θ. Therefore, we have

$$\min_{J\in\mathfrak{J}^1_\theta}\operatorname{Tr} J_\theta^{S}J^{-1} > \max_{J\in\mathfrak{J}^{LOCC,1}_\theta}\operatorname{Tr} J_\theta^{S-1}J$$

unless $\beta_\theta^A = \beta_\theta^B$ and the $+$-sign in (10) is the case. Given $J_\theta^{S,A}$ and $J_\theta^{S,B}$, the gap between the both ends becomes largest when $\beta_\theta^A = \beta_\theta^B = 1$ and $\beta_\theta = 0$.

Example 2 *Let $\left|\phi_\theta^A\right\rangle$ qubit states, and $\left|\phi_\theta^B\right\rangle$ be its complex complement, i.e.,*

$$\left|\phi_\theta^A\right\rangle = \left[\begin{array}{c} e^{-\sqrt{-1}\frac{\theta^2}{2}} \cos\frac{\theta^1}{2} \\ e^{\sqrt{-1}\frac{\theta^2}{2}} \sin\frac{\theta^1}{2} \end{array}\right], \left|\phi_\theta^B\right\rangle = \left[\begin{array}{c} e^{\sqrt{-1}\frac{\theta^2}{2}} \cos\frac{\theta^1}{2} \\ e^{-\sqrt{-1}\frac{\theta^2}{2}} \sin\frac{\theta^1}{2} \end{array}\right],$$

It is easy to check

$$J_\theta^{S,A} = J_\theta^{S,B},$$

and

$$\beta_\theta^A = \beta_\theta^B = 1,$$
$$\beta_\theta = 0.$$

This example is so called 'unti-copy'. With LOCC, one can never make use of the effect of having unit-copy.

A natural generalization of unit-copy would be the state family with $\beta_{\theta,j} = 0$ $(\forall j)$. This is necessary and sufficient [10] for being

$$\mathfrak{J}_\theta^1 = \left\{J \ ; \ J \le J_\theta^S\right\}.$$

Due to (8) and lemma 1, $\mathfrak{J}_\theta^1 = \mathfrak{J}_\theta^{LOCC,1}$ occurs if and only if $\mathfrak{J}_\theta^{A,1} = \left\{J \ ; \ J \le J_\theta^{S,A}\right\}$ and $\mathfrak{J}_\theta^{B,1} = \left\{J \ ; \ J \le J_\theta^{S,B}\right\}$. This condition is equivalent to $\beta_{\theta,j}^A = \beta_{\theta,j}^B = 0$ $(\forall j)$. Hence, if $\beta_{\theta,j} = 0$ $(\forall j)$ and $\beta_{\theta,j}^A \neq 0$ $(\exists j)$, the gap between LOCC an the globally optimal measurement can be observed.

Remark 3 *A tricky case is that $|\phi_\theta\rangle$ is separable only for $\theta \in \Theta_s \subset \Theta$. Suppose Θ is an open set in $\mathbb{R}^{d'}$, and that that the map $\theta \to |\phi_\theta\rangle$ is smooth. Then, $\dim \Theta_s$ has to be strictly smaller than $\dim \Theta$. Therefore, if we consider an arbitrary prior distribution of θ which can be written as $q(\theta)\,\mathrm{d}\theta$, with $\mathrm{d}\theta$ being the Lebesgue measure, it won't contribute to the average of the figure of merit with respect to the prior distribution.*

5.3 Proof of lemma 1

The proof is similar to the proof of lemma 1 in [4]. Let x_t and y_t be the data obtained at tth round, and define $x^{t-1} := x_1 \cdots x_{t-1}$, and $y^{t-1} := y_1 \cdots y_{t-1}$. The measurement by Alice and Bob at tth round is denoted by $A_t^{x^{t-1}y^{t-1}}$ and $B_t^{x^{t-1}y^{t-1}}$, respectively.

The key point is that in optimization (5), the measurement M^n can depend on the true value of θ. Note that this is not the case for the estimation scheme minimizing (3). However, to compute the first order asymptotic term of (3), we only have to solve the optimization problem (5), where the measurement may depend on unknown parameter θ. The fact that θ is unknown is reflected in the fact the Fisher information is a function of the derivative of the probability distribution with respect to θ.

Suppose also Alice has $\rho_{A,\theta}^{\otimes n} \otimes \rho_{B,\theta_0}^{\otimes n}$ and Bob has $\rho_{A,\theta_0}^{\otimes n} \otimes \rho_{B,\theta}^{\otimes n}$, locally. Instead of doing communication, Alice applies $A_t^{x^{t-1}y^{t-1}}$ to $\rho_{A,\theta}^{\otimes n}$, and $B_t^{x^{t-1}y^{t-1}}$ to $\rho_{B,\theta_0}^{\otimes n}$, and Bob also does the same. If $\theta = \theta_0$, Fisher information matrix of this local measurement scheme is the scheme equals the one of the LOCC measurement scheme specified by $A_t^{x^{t-1}y^{t-1}}$ and $B_t^{x^{t-1}y^{t-1}}$, as is shown below. Since the construction of measurement can depend on the value of θ, we have the lemma.

The measurement scheme corresponding to this optimum solution of (5) is as follows. Alice and Bob measures $\sqrt{n}$ copies of ρ_θ locally, and exchange the measurement data. They compute auxiliary estimate $\tilde{\theta}_n$. Believing that this value is true, Alice and Bob fabricate $\rho_{B,\tilde{\theta}_n}^{\otimes n-\sqrt{n}}$ and $\rho_{A,\tilde{\theta}_n}^{\otimes n-\sqrt{n}}$, and applies the LOCC measurement optimal at $\theta = \tilde{\theta}_n$ to $\left(\rho_{A,\theta} \otimes \rho_{B,\tilde{\theta}_n}\right)^{\otimes n-\sqrt{n}}$ and $\left(\rho_{A,\tilde{\theta}_n} \otimes \rho_{B,\theta}\right)^{\otimes n-\sqrt{n}}$, respectively. Finally, they exchange the measurement data, and compute the estimate.

Below, we assume $\dim \theta = 1$ for simplicity, but general case is a trivial generalization. If the LOCC measurement M^n is realized by LOCC with t rounds of exchange of classical communication, denoting

$$p_\theta\left(x^t|y^{t-1}\right) = \Pr_\theta\left\{x^t|y^t\right\}$$
$$q_\theta\left(y^t|x^{t-1}\right) = \Pr_\theta\left\{y^t|x^t\right\},$$

we have

$$\begin{aligned}
&J_\theta^{M^n}\Big|_{\theta=\theta_0} \\
&= \sum_{x^t,y^t}\left[\begin{array}{c} p_\theta\left(x^t|y^{t-1}\right) q_\theta\left(y^t|x^{t-1}\right) \\ \times\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln p_\theta\left(x^t|y^{t-1}\right) q_\theta\left(y^t|x^{t-1}\right)\right)^2 \end{array}\right]_{\theta=\theta_0} \\
&= \sum_{x^t,y^t}\left[\begin{array}{c} q_\theta\left(y^t|x^{t-1}\right) \times \\ p_\theta\left(x^t|y^{t-1}\right)\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln p_\theta\left(x^t|y^{t-1}\right)\right)^2 \end{array}\right]_{\theta=\theta_0} \\
&+ \sum_{x^t,y^t}\left[\begin{array}{c} p_\theta\left(x^t|y^{t-1}\right) \times \\ q_\theta\left(y^t|x^{t-1}\right)\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln q_\theta\left(y^t|x^{t-1}\right)\right)^2 \end{array}\right]_{\theta=\theta_0} \\
&+ \sum_{x^t,y^t}\left[\begin{array}{c} p_\theta\left(x^t|y^{t-1}\right) q_\theta\left(y^t|x^{t-1}\right) \\ \times\frac{\mathrm{d}}{\mathrm{d}\theta}\ln q_\theta\left(y^t|x^{t-1}\right)\frac{\mathrm{d}}{\mathrm{d}\theta}\ln p_\theta\left(x^t|y^{t-1}\right) \end{array}\right]_{\theta=\theta_0}.
\end{aligned}$$

Here, observe the last term equals 0, due to the

following reason. We have

$$
\begin{aligned}
&\sum_{x^t,y^t} p_\theta\left(x^t|y^{t-1}\right) q_\theta\left(y^t|x^{t-1}\right) \\
&\quad\times \frac{\mathrm{d}}{\mathrm{d}\theta}\ln p_\theta\left(x^t|y^{t-1}\right)\frac{\mathrm{d}}{\mathrm{d}\theta}\ln q_\theta\left(y^t|x^{t-1}\right) \\
&=\sum_{x^t,y^t}\left[\begin{array}{c} p_\theta\left(x_t|x^{t-1}y^{t-1}\right) p_\theta\left(x^{t-1}|y^{t-2}\right) \\ \times q_\theta\left(y_t|y^{t-1}x^{t-1}\right) q_\theta\left(y^{t-1}|x^{t-2}\right) \\ \times\left(\begin{array}{c}\frac{\mathrm{d}}{\mathrm{d}\theta}\ln p_\theta\left(x_t|x^{t-1}y^{t-1}\right) \\ +\frac{\mathrm{d}}{\mathrm{d}\theta}\ln p_\theta\left(x^{t-1}|y^{t-2}\right)\end{array}\right) \\ \times\left(\begin{array}{c}\frac{\mathrm{d}}{\mathrm{d}\theta}\ln q_\theta\left(y_t|y^{t-1}x^{t-1}\right) \\ +\frac{\mathrm{d}}{\mathrm{d}\theta}\ln q_\theta\left(y^{t-1}|x^{t-2}\right)\end{array}\right)\end{array}\right] \\
&=\sum_{x^{t-1},y^{t-1}}\left[\begin{array}{c} p_\theta\left(x^{t-1}|y^{t-2}\right) q_\theta\left(y^{t-1}|x^{t-2}\right) \\ \times\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln p_\theta\left(x^{t-1}|y^{t-2}\right)\right) \\ \times\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln q_\theta\left(y^{t-1}|x^{t-2}\right)\right)\end{array}\right] \\
&=\sum_{x_1 y_1} p_\theta\left(x_1\right) q_\theta\left(y_1\right)\frac{\mathrm{d}}{\mathrm{d}\theta}\ln p_\theta\left(x_1\right)\frac{\mathrm{d}}{\mathrm{d}\theta}\ln q_\theta\left(y_1\right) \\
&=0.
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
&J_\theta^{M^n}\Big|_{\theta=\theta_0} \\
&=\sum_{x^t,y^t}\left[\begin{array}{c} q_{\theta_0}\left(y^t|x^{t-1}\right)\times \\ p_\theta\left(x^t|y^{t-1}\right)\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln p_\theta\left(x^t|y^{t-1}\right)\right)^2\end{array}\right]_{\theta=\theta_0} \\
&+\sum_{x^t,y^t}\left[\begin{array}{c} p_{\theta_0}\left(x^t|y^{t-1}\right)\times \\ q_\theta\left(y^t|x^{t-1}\right)\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln q_\theta\left(y^t|x^{t-1}\right)\right)^2\end{array}\right]_{\theta=\theta_0}
\end{aligned}
$$

Therefore, the first term is the average of the Fisher information of the probability distribution family $\left\{p_\theta\left(x^t|y^{t-1}\right)\right\}_{\theta\in\Theta}$ with y^t obeying $q_{\theta_0}\left(y^t|x^{t-1}\right)$. The second term is the similar.

6 Acknowledgment

The author is thankful to M. Hayashi for pointing out that the proof of lemma 1 can be done in the analogous way as the proof of lemma 1 in [4].

References

[1] F. Anselmi, A. Chefles and M. B. Plenio, New. J. Phys. 6, 164 (2004).

[2] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres and W. K. Wootters, " Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," Phys. Rev. Lett. **70**, 1895, (1993).

[3] R. Goodman and N. Wallach, *Representations and Invariants of the Classical Groups*, (Cambridge University Press, 1998.

[4] M. Hayashi, K. Matsumoto, "Two kinds of Bahadur-type bound in adaptive experimental design", IEICE Transaction (Japanese Edition) A, Vol. J83-A No. 6, 629-638 (2002).

[5] M. Hayashi, D. Markham, M. Murao, M. Owari, S. Virmani, "Bounds on Multipartite Entangled Orthogonal State Discrimination Using Local Operations and Classical Communication", Phys. Rev. Lett. 96 ,040501 (2006)

[6] M. Hayashi and K. Matsumoto, "Quantum universal variable-length source coding," Phys. Rev. A **66**, 022311(2002).

[7] M. Owari, M. Hayashi, "Local copying and local discrimination as a study for non-locality of a set", Phys. Rev. A 74, 032108 (2006)

[8] M. Hayashi, K. Matsumoto, and Y. Tsuda, "A study of LOCC-detection of a maximally entangled state using hypothesis testing", J. of Phys. A, 39 14427-14446 (2006).

[9] M. Keyl, R. F. Werner , "Optimal Cloning of Pure States, Judging Single Clones," J.Math.Phys. 40 3283-3299(1999)

[10] K. Matsumoto, "A new approach to the Cramer-Rao-type bound of the pure-state model," Journal of Physics A vol. 35 No. 13 (5 April 2002) 3111-3123.

[11] K. Matsumoto, Entanglement cost and distillable entanglement of symmetric states, quant-ph/0708.3129 (2007).

[12] K. Matsumoto and M. Hayashi, " Universal entanglement concentration", Phys. Rev. A 75, 062338 (2007) quant-ph/0509140.

[13] R.F.Werner, " "Optimal cloning of pure states", Phys.Rev.A58, 1827(1998)

[14] H. Weyl, *The Classical Groups, their invariants and representations*, Princeton (1939).

A Group representation theory

Lemma 4 *Let U_g and U_g' be an irreducible representation of G on the finite-dimensional space $\mathcal{H}$ and $\mathcal{H}'$, respectively. We further assume that U_g and U_g' are not equivalent. If a linear operator A in $\mathcal{H}\oplus\mathcal{H}'$ is invariant by the transform $A\to U_g\oplus U_g' A U_g^*\oplus U_g'^*$ for any g, $\mathcal{H}A\mathcal{H}'=0$ [3].*

Lemma 5 *(Shur's lemma [3]) Let U_g be as defined in lemma 4. If a linear map A in $\mathcal{H}$ is invariant by the transform $A\to U_g A U_g^*$ for any g, $A=c\mathrm{Id}_{\mathcal{H}}$.*

B Representation of symmetric group and SU

Due to [3], we have

$$\dim \mathcal{U}_\lambda = \frac{\prod_{i<j}(l_i - l_j)}{\prod_{i=1}^{d-1}(d-i)!}, \tag{11}$$

$$d_\lambda = \dim \mathcal{V}_\lambda = \frac{n!}{\prod_{i=1}^{d}(\lambda_i + d - i)!}\prod_{i<j}(l_i - l_j), \tag{12}$$

with $l_i := \lambda_i + d - i$. It is easy to show

$$\log \dim \mathcal{U}_\lambda \le d^2 \log n. \tag{13}$$

Let $a_\lambda^\phi = \mathrm{Tr}\left\{\mathcal{W}_{\lambda,A}\left(\mathrm{Tr}_B|\phi\rangle\langle\phi|\right)^{\otimes n}\right\}$ and the formulas in the appendix of [6] says

$$\left|\frac{\log d_\lambda}{n} - \mathrm{H}\left(\frac{\lambda}{n}\right)\right| \le \frac{d^2+2d}{2n}\log(n+d), \tag{14}$$

$$\sum_{\frac{\lambda}{n}\in\mathrm{R}} a_\lambda^\phi \le (n+1)^{d(d+1)/2}\exp\left\{-n\min_{\boldsymbol{q}\in\mathrm{R}}\mathrm{D}\left(\boldsymbol{q}||\boldsymbol{p}\right)\right\}, \tag{15}$$

where R is an arbitrary closed subset.

C